Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **466,** 3331–3343 (2017)
Advance Access publication 2016 December 15

doi:10.1093/mnras/stw3267

Black hole clustering and duty cycles in the Illustris simulation

C. DeGraf★ and D. Sijacki

Institute of Astronomy and Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

Accepted 2016 December 13. Received 2016 December 13; in original form 2016 September 20

ABSTRACT

We use the high-resolution cosmological simulation Illustris to investigate the clustering of supermassive black holes across cosmic time, the link between black hole clustering and host halo masses, and the implications for black hole duty cycles. Our predicted black hole correlation length and bias match the observational data very well across the full redshift range probed. Black hole clustering is strongly luminosity dependent on small, 1-halo scales, with some moderate dependence on larger scales of a few Mpc at intermediate redshifts. We find black hole clustering to evolve only weakly with redshift, initially following the behaviour of their hosts. However, below $z \sim 2$ black hole clustering increases faster than that of their hosts, which leads to a significant overestimate of the clustering-predicted host halo mass. The full distribution of host halo masses is very wide, including a low-mass tail extending up to an order of magnitude below the naive prediction for minimum host mass. Our black hole duty cycles, $f_{\rm duty}$, follow a power-law dependence on black hole mass and decrease with redshift, and we provide accurate analytic fits to these. The increase in clustering amplitude at late times, however, means that duty cycle estimates based on black hole clustering can overestimate $f_{\rm duty}$ substantially, by more than two orders of magnitude. We find the best agreement when the minimum host mass is assumed to be $10^{11.2}\,{\rm M}_{\odot}$, which provides an accurate measure across all redshifts and luminosity ranges probed by our simulation.

Key words: black hole physics – methods: numerical – galaxies: active – galaxies: haloes – quasars: general.

1 INTRODUCTION

It is now widely understood that supermassive black holes are found at the centre of massive galaxies (Kormendy & Richstone 1995), and that properties of the host galaxy strongly correlate with black hole mass (e.g. Magorrian et al. 1998; Gebhardt et al. 2000; Graham et al. 2001; Ferrarese 2002; Tremaine et al. 2002; Häring & Rix 2004; Gültekin et al. 2009; Kormendy & Ho 2013; McConnell & Ma 2013).

One fundamental aspect of black hole studies is clustering behaviour that provides a unique way of linking black holes to their host galaxies. Black hole clustering has been studied extensively in observations (e.g. La Franca, Andreani & Cristiani 1998; Porciani, Magliocchetti & Norberg 2004; Croom, Boyle, Shanks et al. 2005; Myers et al. 2007; Shen et al. 2007; da Ângela et al. 2008; Ross et al. 2009; Shen et al. 2009; White et al. 2012; Eftekharzadeh et al. 2015; Ikeda et al. 2015), as well as simulations (e.g. Bonoli et al. 2009; Croton 2009; DeGraf, Di Matteo & Springel 2010; DeGraf et al. 2012).

Since the emergence of large-scale surveys capable of probing a range of redshifts, the general consensus is that the clustering signal decreases with time. At low redshift (below $z \sim 2$), the evidence for redshift evolution is generally weak, but at higher redshifts the evidence is much stronger, with correlation lengths approaching $10\ h^{-1}$ Mpc (e.g. Myers et al. 2006; White et al. 2012) and bias factors (the clustering strength relative to that of the underlying dark matter density distribution) as large as $b = 5$–10 (e.g. Shen et al. 2009; Ikeda et al. 2015). In addition to the redshift evolution, the possibility of luminosity dependent clustering has crucial implications for our understanding of the relation between black holes and their host haloes. In particular, under the simplistic assumption that active galactic nuclei (AGN) luminosity is proportional to host halo mass, one would expect brighter samples to be more strongly clustered (consistent with the stronger clustering of more massive haloes). On the other hand, most models suggest a more widely varying black hole luminosity history, such that both bright and faint AGN can populate similar haloes at different phases of their lifetimes, in which case clustering behaviour should only weakly depend on instantaneous luminosity. Many observations have found a lack of luminosity dependence (e.g. Croom et al. 2005; Myers et al. 2007; da Ângela et al. 2008; White et al. 2012; Krolewski & Eisenstein 2015) or only a weak dependence (e.g. Shen et al. 2009; Eftekharzadeh et al. 2015), supporting this model. Work by Bonoli et al. (2009) suggests, however, that even in the case of varying luminosity histories, a luminosity dependence could be found among lower luminosity black holes, whose simulations are well suited to investigate as observations being to push to lower flux limits.

★ E-mail: cdegraf@ast.cam.ac.uk

By matching quasar clustering to that of dark matter haloes, the typical mass of the haloes that host quasars can be estimated, providing a relatively simple means of estimating host properties for a range of black hole populations. By taking the expected number density of such haloes and combining with the number density of AGN (via a luminosity function), one can estimate the active fraction of black holes or duty cycle (see, e.g. Haiman & Hui 2001; Martini & Weinberg 2001; Grazian et al. 2004; Shankar et al. 2010). Such duty cycle estimates, however, rely upon several assumptions, most significantly the accuracy of the typical and/or minimum host halo mass calculated from clustering. Furthermore, these estimates rely upon the link between AGN luminosity and the host halo. Some models, however, suggest that only peak AGN luminosity correlates with host halo mass (e.g. Hopkins et al. 2005a,b); in these models the scatter between low-luminosity lifetimes and host properties suggests that clustering should have a weaker luminosity dependence (e.g. Lidz et al. 2006) and that the assumptions used when estimating duty cycles may not hold. Large-scale cosmological simulations are well suited for investigating these aspects of black hole clustering, which we focus on in this paper.

Here, we use the state-of-the-art Illustris simulation (Nelson et al. 2015) to study the clustering of supermassive black holes across cosmic time, taking advantage of the statistically representative sample provided by a large-volume simulation. The Illustris simulation is a $(106.5\ \mathrm{Mpc})^3$ box, providing a sufficiently large sample to predict clustering behaviour in detail for $z = 0$–4, including dependence on black hole luminosity. The Illustris simulation has been shown to reproduce several key black hole properties, including black hole mass density, mass function, luminosity function and scaling relations (Sijacki et al. 2015), making it ideally suited to investigate black hole clustering behaviour. In addition to showing the clustering via the black hole autocorrelation function, we compare the clustering amplitude via both the correlation length and bias parameters to observational results. Using the strength of the black hole clustering signal, we are able to estimate the typical mass of host haloes similar to observational approaches, and directly compare to the actual distribution of host masses and explain the discrepancies therein. Similarly, clustering can be used to estimate the duty cycle of black holes (see, e.g. Eftekharzadeh et al. 2015), which we compare directly to the actual duty cycle in the simulation, probing the accuracy of this estimate for several different definitions for duty cycle.

The outline for our paper is as follows. In Section 2, we outline the numerical methods used, including the Illustris simulation project, and the clustering calculation used throughout the paper. In Section 3, we discuss the results of our investigation. Section 3.1 covers the clustering of black holes, their evolution with redshift and dependence on black hole luminosity. In Section 3.2, we link the clustering behaviour to properties of host haloes. In Section 3.3, we characterize the duty cycle of both black holes and haloes, and the issues involved in estimated duty cycle from clustering behaviour. Finally, we summarize our conclusions in Section 4.

2 METHOD

2.1 Simulations

In this study, we analyse the Illustris[1] suite of simulations performed using the hydrodynamical code AREPO (Springel 2010). This code uses a TREEPM gravity solver and a second-order unsplit Godunov method for solving for the hydro forces. The hydrodynamics equations are solved on an unstructured Voronoi mesh that can move with the fluid in a quasi-Lagrangian manner. Numerous computational and cosmological tests have been performed on the code, verifying its ability to properly capture shock properties, develop fluid instabilities and maintain low numerical diffusivity and Galilean invariance (see, e.g. Springel 2010, 2011; Bauer & Springel 2012; Kereš et al. 2012; Sijacki et al. 2012; Torrey et al. 2012; Vogelsberger et al. 2012; Nelson et al. 2013).

The Illustris suite of simulations (Genel et al. 2014; Vogelsberger et al. 2014a) analysed here uses a $(106.5\mathrm{Mpc})^3$ cosmological box at several resolutions, with dark matter only, non-radiative and full hydrodynamic runs. For this work, we focus only on the full-hydro high-resolution simulation (Illustris-1). This run uses a standard Λ cold dark matter cosmology, with $\Omega_{\mathrm{m},0} = 0.2726$, $\Omega_{\Lambda,0} = 0.7274$, $\Omega_{\mathrm{b},0} = 0.0456$, $\sigma_8 = 0.809$, $n_s = 0.963$, $H_0 = 70.4\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ (consistent with Hinshaw et al. 2013) and runs from $z_{\mathrm{start}} = 127$ to $z_{\mathrm{end}} = 0$. The simulation has 3×1820^3 resolution elements with typical gas cell mass $m_{\mathrm{gas}} = 1.26 \times 10^6\ \mathrm{M}_\odot$ and gravitational softening $\epsilon_{\mathrm{gas}} = 0.71$ kpc and dark matter particle mass $m_{\mathrm{DM}} = 6.26 \times 10^6\ \mathrm{M}_\odot$ and gravitational softening $\epsilon_{\mathrm{DM}} = 1.42$ kpc.

The Illustris simulations include a detailed model of the physics involved in galaxy formation. Primordial and metal-line cooling are included in the presence of a time-dependent ultraviolet background (Faucher-Giguère et al. 2009) including self-shielding (Rahmati et al. 2013); star formation and associated supernovae feedback follow the model of Springel & Hernquist (2003), using a softer equation of state (Springel et al. 2005) with $q = 0.3$ and a Chabrier initial mass function (Chabrier 2003). Models for stellar evolution, gas recycling and metal enrichment are also included (see also Wiersma et al. 2009), along with mass- and metal-loaded galactic outflows (see also Oppenheimer & Davé 2008; Okamoto et al. 2010; Puchwein & Springel 2013).

Black holes are treated as collisionless sink particles. Seeding is based on an on-the-fly friends-of-friends algorithm, where black hole particles with seed mass of $10^5\,h^{-1}\ \mathrm{M}_\odot$ are inserted into haloes with mass above $5 \times 10^{10}\,h^{-1}\ \mathrm{M}_\odot$ that do not already contain a black hole particle. This seeding prescription is intended to remain consistent with a variety of formation models, including direct collapse of gas clouds (e.g. Bromm & Loeb 2003; Begelman, Volonteri & Rees 2006), or formation of smaller seeds from early PopIII stars (Bromm & Larson 2004; Yoshida et al. 2006) followed by a period of rapid growth leading to our seed mass. After seeding, black holes grow though accretion via gas accretion and black hole mergers, which occur when a pair of black holes pass within their respective smoothing lengths. Gas accretion is based on a Bondi–Hoyle-like formalism ($\dot{M}_{\mathrm{BH}} = (4\pi\alpha G^2 M_{\mathrm{BH}}^2 \rho)/c_s^3$, Bondi & Hoyle 1944; Bondi 1952), with an imposed upper limit of the Eddington rate [$\dot{M}_{\mathrm{Edd}} = (4\pi G M_{\mathrm{BH}} m_p)/(\epsilon_r \sigma_T c)$]. In addition, a pressure criterion is applied to lower the accretion rate if the ambient gas pressure is insufficient to compress the gas above the star formation density threshold, preventing formation of unrealistic bubbles in low-density gas (see Vogelsberger et al. 2013, for more details).

Black hole feedback is included in three separate modes: 'quasar', 'radio' and 'radiative'. In quasar mode, feedback is radiated with an efficiency of $\epsilon_r = 0.2$, and couples thermally to the surrounding gas with an efficiency $\epsilon_f = 0.05$. For black holes with accretion efficiency below $\chi_{\mathrm{radio}} = \dot{M}_{\mathrm{BH}}/\dot{M}_{\mathrm{Edd}} = 0.05$, the radio mode is used. Radio mode feedback is applied by inserting energy into hot bubbles randomly placed around the black hole, representing bubbles

[1] http://www.illustris-project.org; Nelson et al. (2015)

expected to be inflated by AGN radio jets, with the bubble energy set by $E_{\rm bubble} = \epsilon_m \epsilon_r \Delta M_{\rm BH} c^2$, with a coupling factor $\epsilon_m = 0.35$. Finally, radiative feedback is incorporated by modifying the photoionization and photoheating rates near accreting black holes. For a more detailed discussion of the black hole accretion and feedback model, see Sijacki et al. (2007) and Sijacki et al. (2015).

We do note several uncertainties regarding black hole modelling as implemented here. First, the seeding mechanism used is based solely on host halo mass, and does not attempt to characterize the physical process behind formation. Different formation pathways can lead to dramatically different initial mass scales (and thus early accretion rates and associated luminosities), ranging from low-mass seeds from PopIII stars to more massive seeds from runaway interactions in dense nuclear star clusters or direct collapse of massive gas clouds. Although the model used here is intended to be consistent with any of these, the different environmental factors associated with each formation mechanism could potentially be imprinted on black hole clustering. However, this effect would be strongest among very low mass black holes and would likely not have a significant impact on the scales considered in this paper (though future work will investigate dependences on black hole seed formation). We also note that the accretion and feedback models used here are limited by the resolution possible in a large volume simulation. In particular, we note that different driving forces behind black hole fuelling could potentially be found in the clustering signal; e.g. the different environments in which one might find merger-induced fuelling (e.g. Di Matteo, Springel & Hernquist 2005) versus instability-driven growth (e.g. Gabor & Bournaud 2014). However, computational limitations currently prevent resolving the formation of such instabilities in sufficient volume to study clustering (though see DeGraf et al. 2016). Investigations into these areas may provide the means to better distinguish mechanisms for formation and growth of supermassive black holes, but are beyond the scope of this paper. For this investigation, we focus on the more general model for black holes, which has been demonstrated to accurately reproduce a wide range of black hole properties and the correlation with host galaxies (see Sijacki et al. 2015).

For further details on the Illustris simulations, see Vogelsberger et al. (2014a,b), Genel et al. (2014) and Sijacki et al. (2015).

2.2 Clustering calculations

To characterize the clustering of black holes in our simulation, we use the black hole autocorrelation function ξ and the correlation length r_0 [the scale at which $\xi(r_0) = 1$]. The correlation function is calculated using the natural estimator

$$\xi(r) = \frac{\rm DD}{\rm RR} - 1 = \frac{\rm DD}{N_{\rm obj}(N_{\rm obj}-1)\frac{\Delta V}{V}} - 1, \quad (1)$$

where DD is the number of data pairs with a separation of $r \pm \Delta r/2$, RR is the number of pairs expected from a random distribution, $N_{\rm obj}$ is the number of objects in the data set, ΔV is the volume of a spherical shell with radius r and thickness Δr and V is the volume of the simulation box. We note that the periodicity of our simulation boundaries means that edge effects are minimal, confirmed by comparisons with the Landy & Szalay (1993) estimator that provides consistent results. For calculation of $\xi_{\rm DM}$ (plotted in Fig. 1 and used in equation 2), we use a sample of ∼400 000 Dark Matter (DM) particles selected at random from the full simulation box, which provides results converged to within a few per cent.

3 RESULTS

3.1 Black hole clustering

In Fig. 1, we show the autocorrelation function of black holes (solid line), dark matter (dashed line) and galaxies ($M_* > 10^{10}\,{\rm M}_\odot$; dot–dashed line) for $z = 0$, 1, 2 and 4, with shaded regions representing the variation across five sequential snapshots (at ∼150 Myr per snapshot). For all but the smallest scales, the intersnapshot variation is minimal, as expected. We also include Poisson error bars for the black hole correlation function, which shows the Poisson errors are even smaller than the intersnapshot variation. The dark matter clustering evolves such that $\xi_{\rm DM}$ increases with time, as predicted by linear growth models (at least at large scales; at smaller scales non-linear growth dominates). The black hole clustering shows much weaker evolution with redshift, such that the bias between $\xi_{\rm BH}$ and $\xi_{\rm DM}$ decreases as we approach lower redshift, which we address more explicitly below.

In addition to the full correlation, we separate $\xi_{\rm BH}$ into 1-halo (dotted blue) and 2-halo (dotted red) components, defined similarly to equation (1): $\xi_{\rm 1h} = {\rm DD_{1h}/RR} - 1$ and $\xi_{\rm 2h} = {\rm DD_{2h}/RR} - 1$, where $\rm DD_{1h}$ ($\rm DD_{2h}$) are pairs of black holes found in the same (different) haloes. The crossover between 1-halo and 2-halo terms occurs at ∼1 Mpc at $z = 0$ and at progressively smaller scales for higher redshift, consistent with the expectation that the typical haloes hosting black holes are larger at lower redshift. We also note that the 2-halo term completely dominates at larger scales; for this reason we use scales above ∼2 Mpc when calculating the best-fitting functions for correlation length and bias factors (see Section 3.2).

To investigate the luminosity dependence of clustering behaviour, in Fig. 2 we show how progressively higher cuts on $L_{\rm BH} = \epsilon_r \dot{M} c^2$ affect $\xi_{\rm BH}$. To more clearly show the effect, we plot the ratio between the correlation function of luminosity-selected samples and that of the full black hole population ($\xi({\rm L_{BH}} > {\rm L_{cut}})/\xi_{\rm BH}$). We note that the luminosity dependence is strongest at intermediate redshift ($z \sim 1$–2), when AGN activity is quite high and before self-regulation has slowed black hole growth. We also see that luminosity dependence is strongest at the smallest scales (well within the 1-halo term). This suggests that more luminous black holes tend to be strongly clustered within individual haloes (and thus strengthening the 1-halo term), which may be explained by more luminous AGN tending to be found in more massive haloes, which in turn tend to host the largest number of satellite black holes necessary to produce a small-scale 1-halo signal (see Degraf et al. 2011b; Chatterjee et al. 2012).

To better characterize the evolution of the large-scale black hole clustering, we use the correlation length (r_0), defined as the scale at which $\xi(r_0) = 1$. In Fig. 3, we show the correlation length for four luminosity-selected black holes samples. We calculate r_0 using a power-law fit to ξ in the range 2–10 Mpc (where the 2-halo term dominates), but note that the value is not sensitive to the exact fitting range selected. Solid lines show black holes selected regardless of mass, while dashed lines only include black holes with $M_{\rm BH} > 10^7\,{\rm M}_\odot$. For the most luminous black holes in our sample ($L_{\rm BH} > 10^{44}$ erg s^{-1}) we find roughly constant r_0, with only a slight increase at $z \sim 1$–1.5 (and minimal change when imposing the cut on $M_{\rm BH}$, since most luminous AGN are above $10^7\,{\rm M}_\odot$). For fainter luminosity cuts, we find that the correlation length tends to decrease with time to a minimum at $z \sim 2$, followed

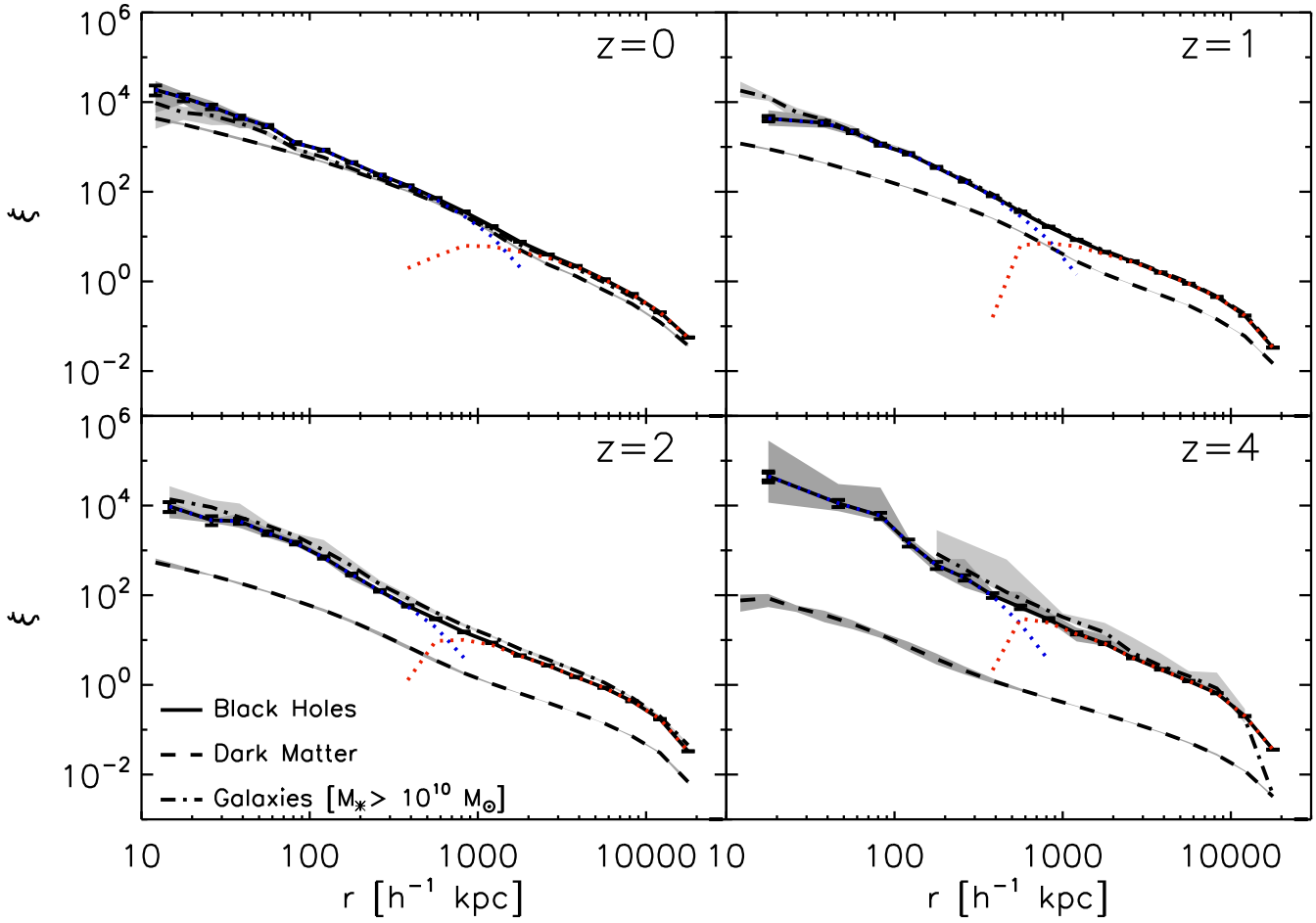

Figure 1. Autocorrelation for black holes (solid line; no luminosity cut and with Poisson error bars), dark matter (dashed line) and galaxies ($M_* > 10^{10}\,M_\odot$; dot–dashed line) at $z = 0$, 1, 2 and 4. The shaded region represents the variation across five sequential snapshots (at ∼150 Myr per snapshot), characterizing the uncertainty in the relation. The blue and red dotted lines are the 1-halo and 2-halo components of the black hole correlation function. The black hole correlation function evolves significantly slower than the dark matter, such that the bias between them decreases with redshift.

by an increase at later times, and fainter black holes tend to be less-strongly clustered. Similar to the semi-analytic work by Bonoli et al. (2009), we find a significant luminosity dependence when considering a sufficiently large range of luminosities. We also show the correlation length of dark matter haloes with $M_{\rm DM} > 10^{11.2}\,M_\odot$, which closely matches the low-luminosity black hole clustering for $z > 2$, suggesting the occupation distribution remains roughly constant at high redshift, consistent with earlier findings (DeGraf et al. 2012). At lower redshifts, the increasing black hole correlation length suggests an increase in typical host halo mass, which we investigate in more detail in Section 3.2.

We also compare with observational measurements of r_0, and find generally consistent measurements. We note that our r_0 values are slightly lower than the observations, however this is at least in part due to the difference between the flux-limited surveys and our volume limited sample. As an example, we show the observations from Eftekharzadeh et al. (2015) adjusted for luminosity (using their luminosity fit) as coloured circles. Although our predictions are still lower than theirs, they are within the error bars, and closely match when considering only black holes with $M_{\rm BH} > 10^{7}\,M_\odot$. We also provide green data points, representing observational data adjusted to match the luminosity of our $L_{\rm BH} > 10^{43}$ erg s^{-1} sample. This adjustment compares the mean luminosity of the observed samples[2] and our z- and $L_{\rm BH}$-selected samples, and applies the luminosity dependence formula from Eftekharzadeh et al. (2015), producing good agreement between our predictions and observed data. We do note that the $L_{\rm BH}$ dependence of Eftekharzadeh et al. (2015) is highly uncertain, however the $L_{\rm BH}$-evolution is consistent with our simulation (see coloured open circles in Fig. 3).

Having shown that the Illustris simulations produce expected behaviour for the black hole autocorrelation functions, we investigate how the clustering behaviour links black hole properties to those of their host haloes. In the bottom panel of Fig. 3, we plot the evolution of the black hole bias, defined as

$$b = \sqrt{\xi_{\rm BH}/\xi_{\rm DM}}\,. \tag{2}$$

We calculate the bias over the range 2–10 Mpc, which keeps us above the 1-halo regime (where non-linear bias can be significant; see Fig. 1 and Degraf, Di Matteo & Springel 2011a), and below the scale at which box-size will be a limiting factor. Observations from a range of studies have been included (see caption), having been adjusted to account for assumed σ_8 value. At low redshift

[2] Where necessary, we apply bolometric corrections based on the quasar Spectral Energy Distribution of Hopkins, Richards & Hernquist (2007).

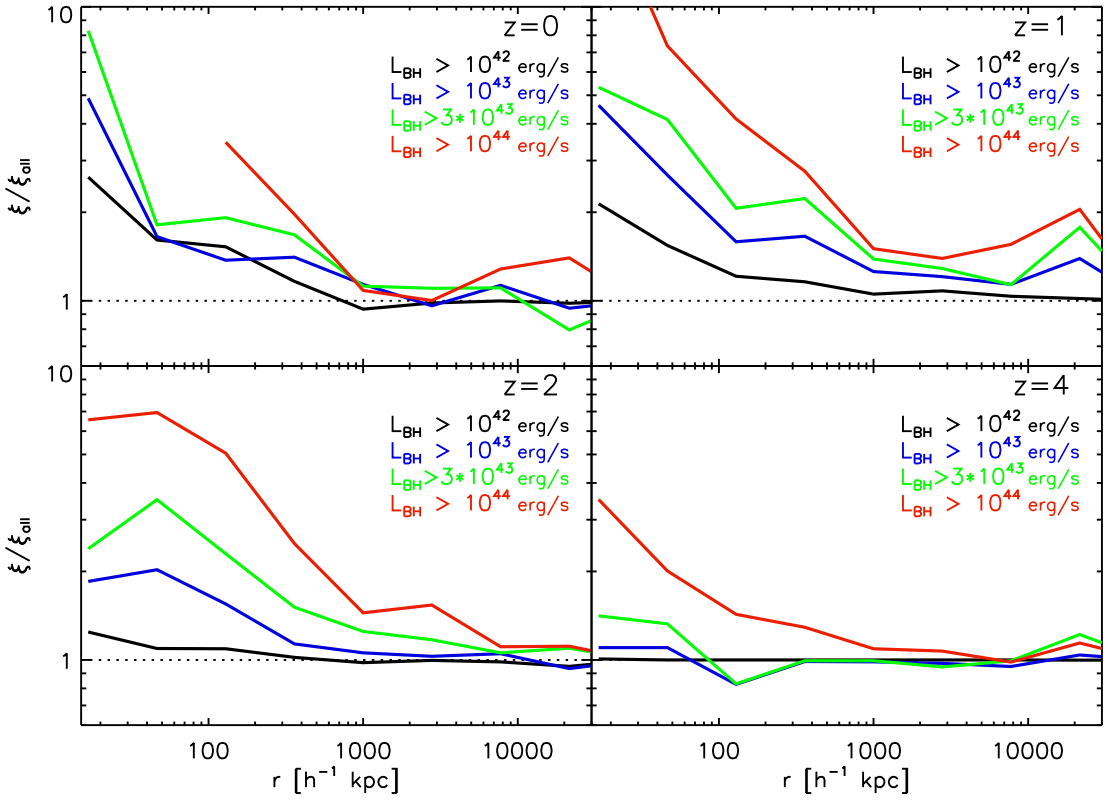

Figure 2. Ratio between autocorrelation functions using luminosity-selected population ξ and full population $\xi_{\rm all}$. For each curve, we take the median ratio between three sequential snapshots (∼150 Myr per snapshot) to smooth out short-term variations. We find that the luminosity-dependent clustering is primarily at small, 1-halo scales, but does extend out to larger scales, especially at intermediate redshift.

(below $z \sim 2$), we have excellent agreement with observations. At higher redshift, we appear to underestimate the bias. However, we note that the disagreement is primarily from the Shen et al. (2009) results (as the Ikeda et al. 2015, results are upper limits only). However, Eftekharzadeh et al. (2015) note that the Shen et al. (2009) measurement is dominated by a single bin at $\sim 35\,h^{-1}$ Mpc, a scale well above that probed by our simulation. We find that the evolution of the black hole bias parameter as a function of redshift is well fit by a second-order polynomial $b(z) = A + Bz + Cz^2$. We provide the best-fitting parameter values in Table 1.

3.2 Host halo properties

In the top panel of Fig. 4, we estimate the typical mass of haloes hosting the luminosity-selected black hole samples. The solid lines show the actual host masses $\langle \log(M_{\rm host}) \rangle$, where we find the expected trend of increasing host mass with time. The dashed lines show the predicted halo mass based solely on the black hole clustering, by matching black hole clustering to that of dark matter haloes. To make this prediction, we calculate the clustering bias over 2–10 Mpc (as in Section 3.1). We use the 2–10 Mpc range to remain in the 2-halo-dominated regime (as seen in Fig. 1). Although limited to 2-halo scales, we note that satellite black holes may none the less contribute to the bias calculation, which we discuss below. We compare this to the halo bias from the formalism of Tinker et al. (2010, see equation 6) adopting a linear matter variance ($\sigma(M)$) calculated using the power spectrum from CAMB[3] (with the cosmological parameters used in the Illustris simulations).[4] We find typical host haloes of $\sim 10^{12}$–$10^{13}\,M_\odot$, consistent with general observational estimates of a few times $10^{12}\,h^{-1}\,M_\odot$ (e.g. Myers et al. 2007; da Ângela et al. 2008; Ross et al. 2009; White et al. 2012; Ikeda et al. 2015).

At high redshift (above $z \sim 2$), the agreement between the two actual mean host mass and the bias-predicted mass is good; for $z < 2$, however, the bias prediction overestimates the typical host mass by a factor of ∼2. At least part of this is due to haloes hosting multiple black holes: larger haloes tend to host larger numbers of satellite black holes, which biases ξ towards the clustering of the larger (and thus more strongly clustered) haloes. We show this explicitly in Fig. 5, where we plot the mean number of black holes per halo as a function of halo mass, showing that massive haloes tend to host multiple black holes above a given $L_{\rm BH,\,min}$. For comparison, the dashed lines in Fig. 5 show the mean number of *central* black holes (defined as the most massive black hole in a given friends-of-friends-defined halo) above $L_{\rm BH,\,min}$, which avoids this issue (note that each halo only has a single central black hole, so this curve has an imposed upper bound of $\langle N_{\rm BH,\,cen} \rangle \leq 1$). At $z = 0$ in particular, we note that $\langle N_{\rm BH,\,cen} \rangle$ actually decreases in the highest mass haloes even as $\langle N_{\rm BH} \rangle$ increases, telling us that in the most massive haloes, the central black hole is often not the most luminous; instead the central black hole has been quenched, while satellite black holes continue to grow more efficiently. We have also used vertical lines

[3] http://camb.info

[4] We have compared the halo correlation function from Illustris to the prediction for equal-mass haloes using this approach, and confirmed excellent agreement.

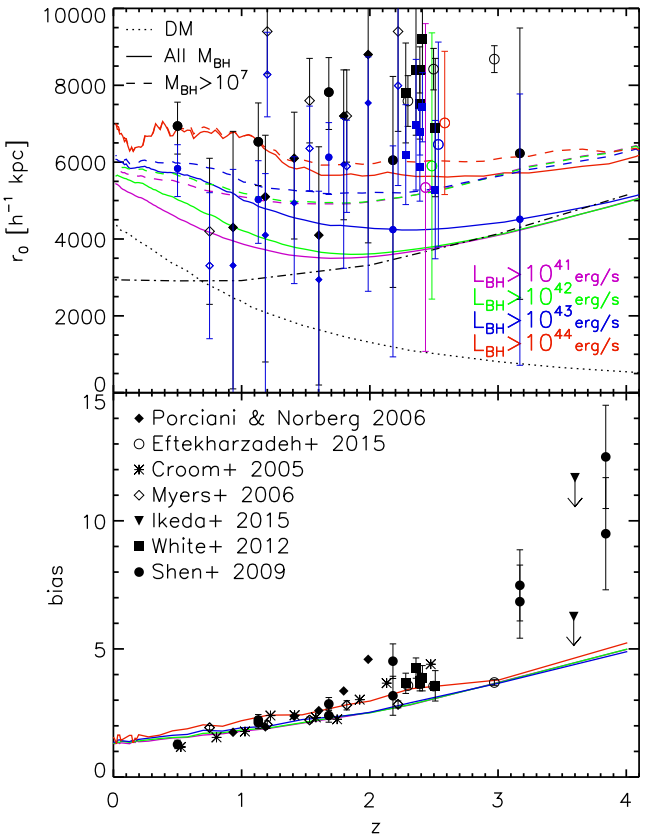

Figure 3. Top: correlation length of black hole autocorrelation function for several luminosity cuts. Curves are smoothed to show overall trend rather than short-term variations in the high-luminosity curve. Solid lines: all black holes included. Dashed lines: only black holes with $M_{\rm BH} > 10^7\,{\rm M}_\odot$ included. Dotted lines: dark matter correlation function. Dot–dashed line: halo correlation function for haloes with $M_{\rm DM} > 10^{11.2}\,{\rm M}_\odot$. The blue coloured points represent observational data that have been adjusted according to the luminosity dependence of Eftekharzadeh et al. (2015) to match the mean luminosity of our $L_{\rm BH} > 10^{43}$ erg s^{-1} sample. We find the correlation length to be moderately dependent on $L_{\rm BH}$, with generally weak evolution with redshift. Bottom: bias (relative to DM autocorrelation) for luminosity selected black holes, compared to observations. Observational data is from Croom et al. (2005), Porciani & Norberg (2006), Myers et al. (2006), Shen et al. (2009), White et al. (2012), Ikeda et al. (2015) and Eftekharzadeh et al. (2015), with bias values adjusted to account for differences in σ_8.

Table 1. Best-fitting parameters for evolution of bias $b(z) = A + Bz + Cz^3$.

$L_{\rm BH,\,min}$	A	B	C
10^{42}	1.343	0.315	0.153
10^{43}	1.390	0.365	0.131
10^{44}	1.468	0.605	0.083

to mark the host mass above which 90 per cent, 75 per cent and 50 per cent of black holes are found. We note that, due to the slope of the halo mass function, most black holes are found in haloes small enough to have $\langle N_{\rm BH}\rangle < 1$ but which are much more common than the more massive haloes.

In the middle panel of Fig. 4, we compensate for this by only considering the central black hole in any given halo, which improves the agreement though some discrepancy remains. To show this more explicitly, in the bottom panel we show the ratio of bias-predicted

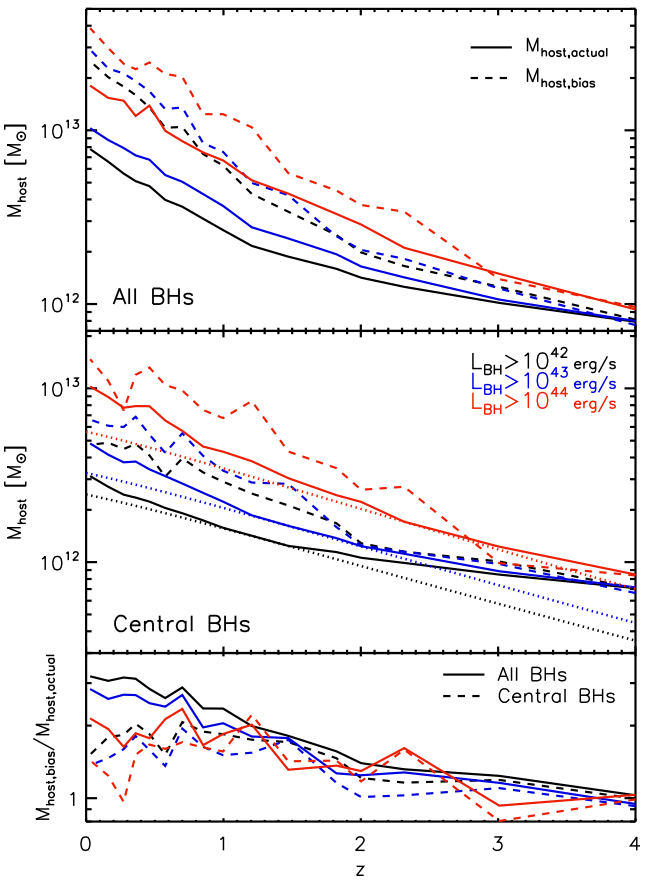

Figure 4. Top: mass of haloes hosting black holes. Solid lines show the actual mean mass (in log-space); dashed lines show the predicted host mass based on the calculated bias. Middle: same as top, but using only central black holes. Dotted lines represent expected halo mass for the median halo growth fitting function of Fakhouri, Ma & Boylan-Kolchin (2010), matching at the redshift where our mean halo mass begins increasing faster than the expected growth curve. Bottom: ratio between bias-predicted host mass and actual host mass, for all black holes (solid lines) and for central black holes only (dashed lines). Black hole clustering increases faster than a $M_{\rm host}$-selected halo sample, leading to an overestimate in the clustering-predicted host halo mass.

mass to the actual host mass for the full black hole population (solid lines) and for the central black holes only (dashed lines). At both high redshifts and high luminosities, the full and central populations agree with one another, since high-redshift and high-luminosity black holes tend to be centrals. At low redshift and moderate to low luminosities, satellite black holes play a larger role and so considering only central black holes decreases the discrepancy by up to a factor of 2, but does not remove it entirely (discrepancies up to $M_{\rm host,\,bias}/M_{\rm host,\,actual} \sim 2$). This suggests that haloes hosting massive, luminous black holes tend to be more strongly clustered than an equivalent halo-mass-selected sample. We test this explicitly in Fig. 6, where we show the halo correlation length in two mass bins, separated into those hosting luminous (blue lines) and faint (red lines) black holes. For each host mass bin, we select the 25 per cent of haloes hosting the most luminous black holes for our bright sample, and the 25 per cent with the least luminous black holes for our faint sample. Here, we clearly see that for a given mass range haloes hosting brighter AGN tend to cluster more strongly than those with faint AGN. In fact, we find that the 10^{11}–$10^{12}\,{\rm M}_\odot$ haloes with the brightest AGN are as strongly clustered as the 10^{12}–$10^{13}\,{\rm M}_\odot$ with

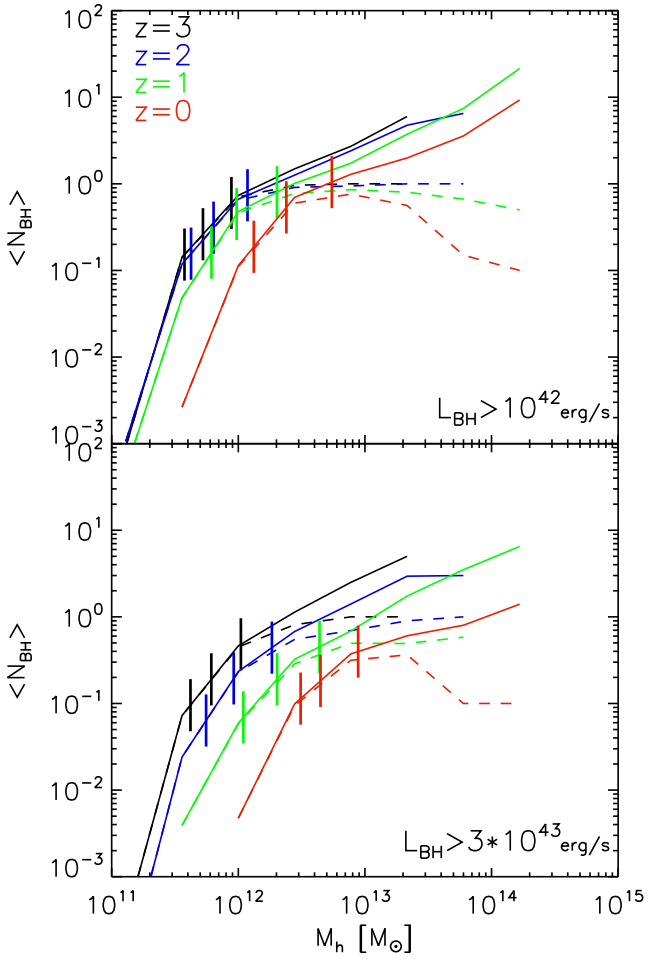

Figure 5. Mean occupation number of black holes above 10^{42} erg s^{-1} (top) and 3×10^{43} erg s^{-1} (bottom). Solid lines show full black hole population; dashed lines show central black holes only. Vertical lines mark the host mass scale above which 90 per cent, 75 per cent and 50 per cent of black holes are found. Despite the lower occupation number, most black holes are found in the more common low-mass haloes, and high-mass haloes ($M_h \gtrsim 10^{12.5}$–10^{13} $M_\odot$) tend to have significant numbers of satellite black holes.

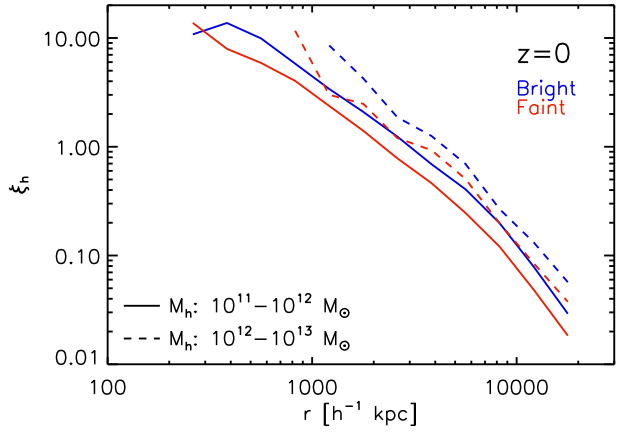

Figure 6. Halo correlation function for haloes hosting bright (blue lines) and faint (red lines) AGN in halo mass ranges of 10^{11}–10^{12} $M_\odot$ (solid lines) and 10^{12}–10^{13} $M_\odot$ (dashed lines). For each halo mass bin, we select the top and bottom quartile in AGN luminosity to form our bright and faint subsamples. This shows that haloes hosting bright AGN tend to be more strongly clustered than those with only faint AGN.

the faintest AGN, despite being significantly smaller (note that the extension of ξ_h to smaller scales characterizes the radial extent of the haloes, in addition to the mass ranges selected). This suggests that observational estimates for typical host halo masses may overestimate by a factor of ∼2, especially at intermediate redshifts where quasars tend to be most active.

We also consider host growth rates in Fig. 4 by adding an expected halo growth curve (dotted lines) according to the median halo growth rate of Fakhouri et al. (2010), matched to the redshift at which the growth of $\langle \log(M_{host}) \rangle$ begins growing faster than the expected median growth rate. We note that our typical host mass does not evolve as a typically growing halo; instead the growth is significantly slower at high redshift, and faster at low redshift. At high redshift, this is largely due to recently seeded black holes: although individual haloes hosting black holes are growing, continued seeding of black holes means that new, low-mass haloes are being added, partially compensating for the growth of the older black holes. At lower redshift, however, a larger black hole population combined with rarer black hole seeding minimizes this effect.

At low redshift, we find that the typical host halo increases faster than the expected median halo growth rate. Here, typical black hole luminosity decreases with time (Sijacki et al. 2015); thus for a given luminosity threshold, as time passes only the most extreme objects continue to satisfy $L_{BH} > L_{BH,\,min}$. In other words, at low redshifts typical L_{BH} decreases, and so the smaller haloes no longer satisfy the luminosity criterion, leading to a faster rise in $\langle M_{host} \rangle$ than the typical halo actually grows.

To more fully characterize typical host haloes, in Fig. 7 we plot the distribution of halo masses hosting black holes above $L_{BH} > 10^{42}$ and 3×10^{43} at $z = 0$ and 4. The solid histograms show the number of black holes above the given luminosity at each halo mass, while the dotted histogram shows the number of central black holes. Vertical lines show the mean halo mass (dotted black line) and the typical mass predicted by the black hole bias parameter (dashed blue line). Consistent with Fig. 4, we see that at high redshift the bias-predicted mass closely matches the actual mean mass. At low redshift, we note that the distribution of host masses increases significantly, and satellite black holes have a strong impact on the high-mass end of the distribution.

One of the calculations sometimes made when interpreting observational data is to estimate the minimum mass of a quasar-hosting halo by considering the mass-averaged halo bias:

$$b_{BH}(>L_{BH,min}) = b_h(M_h > M_{h,min}) = \frac{\int_{M_{h,min}}^{\infty} b(M) \frac{dn}{dM} dM}{\int_{M_{h,min}}^{\infty} \frac{dn}{dM} dM}. \quad (3)$$

However, this only holds if a black hole above $L_{BH,\,min}$ is equally likely to be found in any halo above $M_{h,\,min}$, which Fig. 5 has shown to be inaccurate. To further characterize the validity of this, we overplot $M_{h,\,min}$ estimated from equation (3) as a dashed red line in Fig. 7. At high redshift, this method slightly overpredicts the minimum halo mass, but is relatively close. At low redshift, however, this substantially overestimates the number of luminous black holes in low-mass haloes: although the fraction of low-mass haloes hosting luminous black holes is low, the slope of the halo mass function is such that a large fraction of black holes above a given $L_{BH,\,min}$ are none the less found in relatively low-mass haloes.

Fig. 8 shows the evolution of the minimum mass calculated based on the bias parameter as in equation (3) [solid lines]. We also plot the halo mass above which haloes are found to host 99 per cent of black holes (dashed lines) and 50 per cent of black holes (dotted lines).

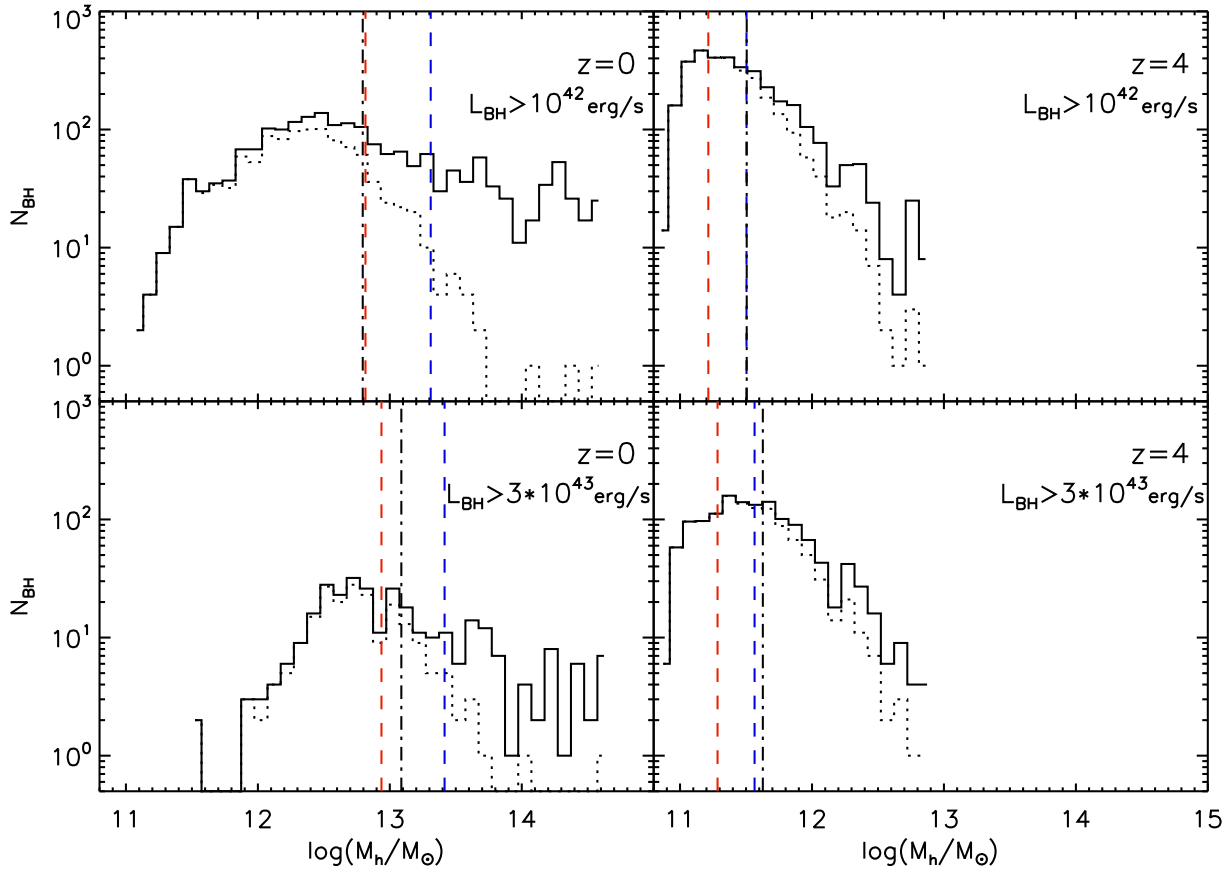

Figure 7. Distribution of host halo masses for black holes above 10^{42} erg s^{-1} (top) and 3×10^{43} erg s^{-1} (bottom) at $z = 0$ (left) and $z = 4$ (right) for full black hole population (solid histogram) and central black hole population (dotted histogram). Dot–dashed black line shows mean host mass. Dashed blue shows the predicted mass based on halo clustering. Dashed red shows the predicted minimum mass based on the bias parameter. The distribution of host halo masses extends below the predicted minimum mass (see equation 3), with a more significant low-end tail at low redshift.

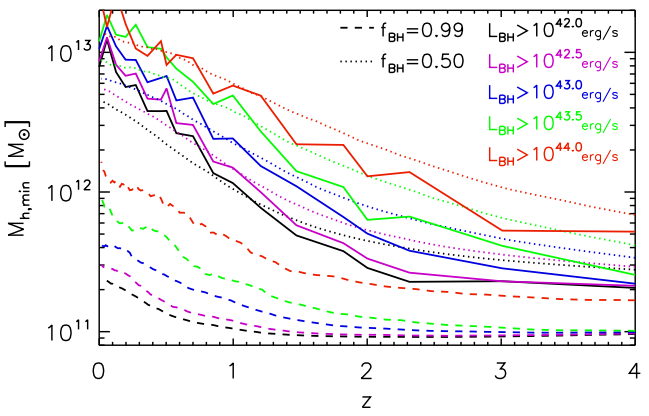

Figure 8. Minimum host mass for several luminosity thresholds calculated using black hole bias (solid lines) and the minimum halo mass above which 99 per cent (dashed lines) and 50 per cent (dotted lines) of black holes are found.

Similar to the curves in Fig. 4, we find that the minimum host mass predicted based on the clustering bias significantly overestimates the actual minimum mass, and is in fact closer to the median mass of host haloes, rather than the minimum. As discussed earlier, this is due to the AGN clustering more strongly than typical haloes of equivalent masses.

3.3 Duty cycle

In addition to the clustering properties, we also consider the duty cycle of black holes in the simulation, and the problems with using clustering behaviour to estimate it. Rather than calculating the fraction of time a given black hole spends above a given luminosity cut, we instead use

$$f_{\rm duty} = \frac{N_{\rm obj}(L_{\rm BH,min} > L_{\rm cut})}{N_{\rm obj}}, \tag{4}$$

which provides us with two main advantages: first, it is based on a single snapshot and thus tracking black holes through mergers does not complicate behaviour, and secondly we can use the same approach to determine both black hole and halo duty cycles (i.e. $N_{\rm obj} = N_{\rm BH}$ and $N_{\rm halo}$).

The black hole duty cycle is plotted in Fig. 9 at $z = 0$, 2 and 4, showing the dependence on both black hole mass and luminosity. In addition to increasing for lower $L_{\rm BH,\,min}$ [a necessary result of equation (4)], more massive black holes tend to have higher duty cycles, as more massive black holes typically have higher accretion rates. At $z = 0$, we find the duty cycle to behave very regularly, with difference $L_{\rm BH,\,min}$ and $M_{\rm BH,\,min}$ thresholds tending to only change the normalization of the duty cycle curve, which are well fit by a power law in $M_{\rm BH,\,min}$.

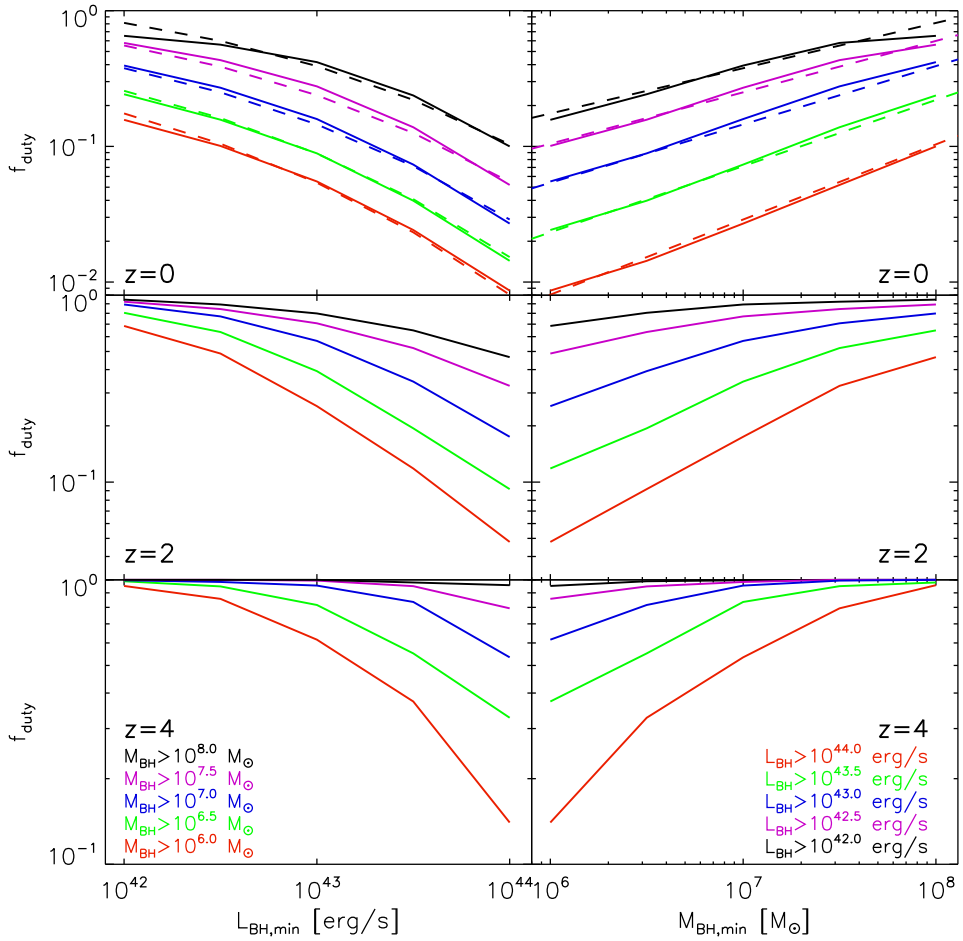

Figure 9. Black hole duty cycle for black holes above a specified minimum mass as a function of $L_{\rm BH}$ (left-hand panels) and specified minimum luminosity as a function of $M_{\rm BH}$ (right-hand panels), for $z = 0$, 2 and 4 (top, middle, bottom). The $z = 0$ panels also include the best-fitting relation from equations (5) and (6). Higher redshift fits are not provided, as the duty cycle rapidly approaches the upper limit of $f_{\rm duty} = 1$, and thus diverges from a well-fit power-law fit (see text for more details). Note that the y-axis range evolves with redshift, matching the range spanned by our simulation.

We have fitted the black hole duty cycle with a power-law form of

$$f_{\rm duty} = 0.1 \times \left(\frac{M_{\rm BH}}{M_0\left(L_{\rm BH,min}\right)} \right)^{\alpha\left(L_{\rm BH,min}\right)}, \quad (5)$$

where M_0 gives us the typical black hole mass at which the duty cycle is 10 per cent, and α characterizes the sensitivity of $f_{\rm duty}$ on $M_{\rm BH}$. Both M_0 and α are dependent on the cut used for $L_{\rm BH}$, and are also well fit by power laws:

$$X = A_X \times \left(\frac{L_{\rm BH,min}}{10^{43}{\rm erg\ s^{-1}}} \right)^{\beta_X}, \quad (6)$$

with $A_{M_0} = 4.2 \times 10^6$, $\beta_{M_0} = 1.35$, $A_\alpha = 0.433$ and $\beta_\alpha = 0.111$ at $z = 0$. We have overplotted these fits as dashed lines in the top panels of Fig. 9, showing excellent agreement. We emphasize that the dashed lines are a single fit over both $M_{\rm BH}$ and $L_{\rm BH}$ using equations (5) and (6), rather than individual fits for each curve. The only discrepancy occurs when the duty cycle increases to above ∼50 per cent, where $f_{\rm duty}$ diverges from a power law as it approaches the maximum possible value of $f_{\rm duty} = 1$. The divergence from a power law is more apparent at higher redshifts: $f_{\rm duty}$ versus $M_{\rm BH}$ is still a rough power law for $f_{\rm duty}$ below ∼0.5, but the higher accretion rates at high z (see Sijacki et al. 2015) produce high duty cycles across all scales. For this reason, we caution that our fitting function for the duty cycle should only be used below $f_{\rm duty} \sim 0.5$ (and at $z = 0$), but this covers the regime of interest when studying duty cycles.

In Fig. 10, we plot the halo duty cycle, rather than the black hole duty cycle, i.e. $N_{\rm obj} = N_{\rm halo}$ in equation (4). In this case, $N_{\rm halo}(L_{\rm BH,\,min})$ is the number of haloes that host at least one black hole above $L_{\rm BH,\,min}$, so $f_{\rm duty}$ provides us with a fraction of haloes that are present above a specified AGN luminosity. As expected, duty cycle increases with halo mass, though we note that the relation tends to flatten out with $M_{\rm halo}$ more rapidly than with $M_{\rm BH}$.

We compare different methods of calculating duty cycle across cosmic time in Fig. 11. Solid lines show the black hole duty cycle (as in Fig. 9) for a range of lower limits on $L_{\rm BH}$, showing the expected decrease in $f_{\rm duty}$ with time. Dotted lines show the halo duty cycle (as in Fig. 10) for haloes above $10^{11.2}$ M$_\odot$, representing a characteristic minimum mass for black hole occupation in our simulation. The $10^{11.2}$ M$_\odot$ threshold was selected to provide a close fit to the black hole duty cycle; increasing the halo mass threshold increases the duty cycle, as seen in Fig. 10. We also consider the duty cycle based on the clustering properties, using the equation

$$f_{\rm duty} = \frac{\int_{L_{\rm BH,min}}^{\infty} \Phi(L){\rm d}L}{\int_{M_{\rm h,min}}^{\infty} \frac{{\rm d}n}{{\rm d}M}{\rm d}M} = \frac{N_{\rm BH}(L_{\rm BH} > L_{\rm BH,min})}{N_{\rm h}(M_{\rm h} > M_{\rm h,min})}, \quad (7)$$

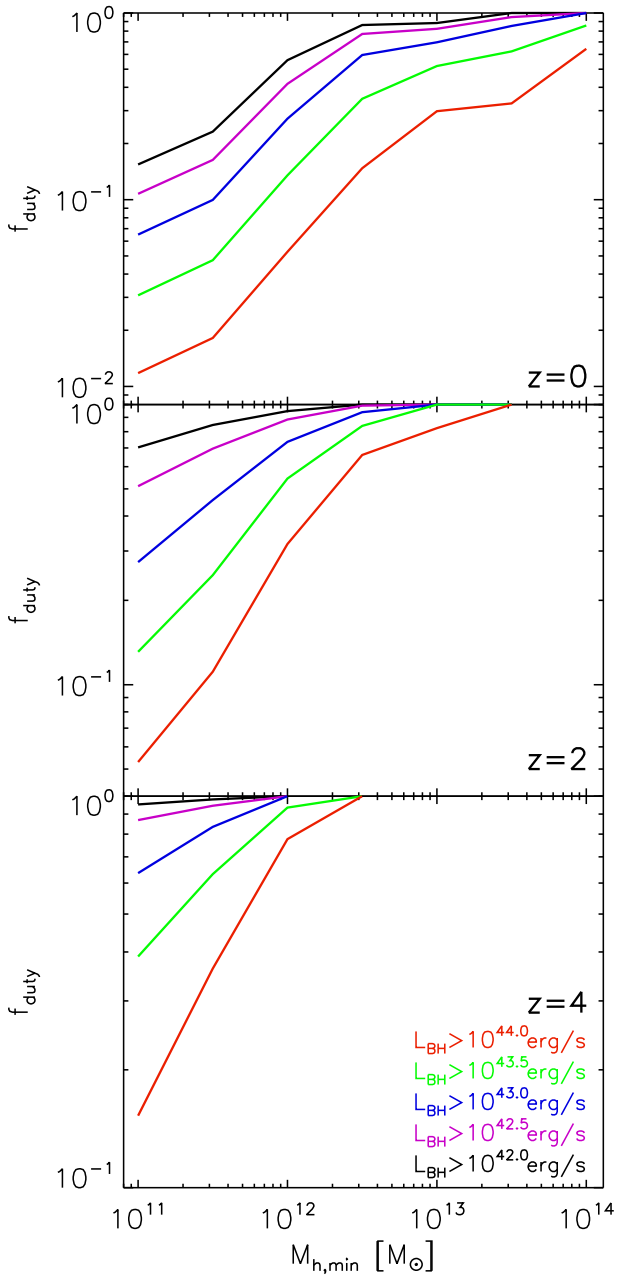

Figure 10. Duty cycle for haloes above a given luminosity (line colour) as a function of minimum halo mass, for $z = 0$, 2 and 4 (top, middle, bottom). Note that the y-axis range evolves with redshift, matching the range spanned by our simulation.

(see, e.g. Martini & Weinberg 2001; Eftekharzadeh et al. 2015), which implicitly assumes that all AGN above $L_{\rm BH,\,min}$ are found in haloes above $M_{\rm h,\,min}$, but no other dependency on halo mass. $L_{\rm BH,\,min}$ is the threshold luminosity used for black holes selection; $M_{\rm h,\,min}$ is the minimum halo mass considered, $\Phi(L)$ is the AGN luminosity function and $\frac{{\rm d}n}{{\rm d}M}$ is the halo mass function. Of particular importance is $M_{\rm h,\,min}$, which is calculated based on the clustering bias as in equation (3). An overestimate of the clustering amplitude would thus produce a correspondingly overestimated $M_{\rm h,\,min}$, and therefore also overestimate the duty cycle. As shown in Fig. 5, the mean occupation number evolves significantly with halo mass contrary to this assumption, suggesting a significant bias between the predicted duty cycle from equation (7) and the 'true' duty cycle.

We plot the estimate from equation (7) using the full black hole sample as dot–dashed lines in the left-hand panel of Fig. 11. We note two main issues here: the duty cycle is generally larger than 1, and the redshift evolution is significantly different from the 'true' black hole duty cycle.

The main factor contributing to $f_{\rm duty} > 1$ is that individual haloes can host multiple AGN above $L_{\rm BH,\,min}$, as shown in Fig. 5, while equation (7) assumes a maximum of one per halo. We account for this in the right-hand panel of Fig. 11 by only including central black holes, which decreases $f_{\rm duty}$ but still has $f_{\rm duty} > 1$, due to the mis-estimate of $M_{\rm h,\,min}$. As shown in Fig. 8, calculating $M_{\rm h,\,min}$ from the black hole bias overpredicts the minimum halo mass. This overestimate means numerous haloes are neglected in the denominator of equation (7), which overestimates the duty cycle. By re-calculating $M_{\rm h,\,min}$ such that 99 per cent of black holes are included (as discussed in Section 3.2 and Fig. 8), we get the dashed line in Fig. 11. This corrected calculation shows a more reasonable duty cycle, but which still evolves very differently from the actual black hole duty cycle (solid lines).

We also fit the redshift evolution of the duty cycle of $M_{\rm BH} > 10^6\,{\rm M}_\odot$ black holes to a logistic function

$$f_{\rm duty} = \frac{1}{1 + {\rm e}^{-k(z-z_0)}}, \tag{8}$$

where k and z_0 are both found to be well fit by power-law functions in $L_{\rm BH,\,min}$, as in equation (6), with $A_k = 0.87$, $\beta_k = -0.127$, $A_{z_0} = 3.13$ and $\beta_{z_0} = 0.338$. The top panel of Fig. 12 shows this fit as dashed lines, demonstrating that the fits are excellent across a full range of redshifts and luminosities. Although the fitting was performed over the range $0 < z < 4$, we have extended the plotting range of Fig. 12 to higher redshift. For $L_{\rm BH,\,min} < 10^{44}$ erg s^{-1}, the evolution out to higher redshift remains excellent. Above $z \sim 5$, there are very few black holes that have grown large enough to reach 10^{44} erg s^{-1}, due to the imposed Eddington limit of the simulation (see Section 2.1). Thus, we expect the duty cycle for the highest luminosity to drop at high redshift, simply due to the lack of sufficiently massive black holes, and the fitting function should not be used. We plot the $L_{\rm BH} > 10^{44}$ erg s^{-1} curve as a dotted line for redshifts at which the fraction of black holes capable of reaching 10^{44} erg s^{-1} at maximum (i.e. Eddington) accretion is less than the predicted duty cycle (dashed line). To confirm this explanation, the bottom panel shows the duty cycle for black hole populations selected by Eddington fraction rather than luminosity, finding the expected smooth increase with redshift. Here, we show that the majority of black holes do approach the Eddington limit for $z > 5$, and so the decrease in the $L > 10^{44}$ erg s^{-1} curve in the upper panel is indeed due to limitations on the mass function at high redshift rather than a change in active fraction.

3.4 Luminosity-dependent cross-correlation with satellite galaxies

One final aspect of our analysis is to use AGN clustering to look for signals of AGN-induced galaxy quenching. Rather than using the black hole autocorrelation used in the rest of our work, here we use the cross-correlation function between black holes and galaxies, with a particular emphasis on satellite galaxies. In particular, if AGN were capable of quenching star formation in satellite galaxies, we would expect to find quenched galaxies to be preferentially found near massive (or strongly accreting) black holes, which should be detectable in the black hole–galaxy cross-correlation function. To investigate this, we used many different selection criteria for both

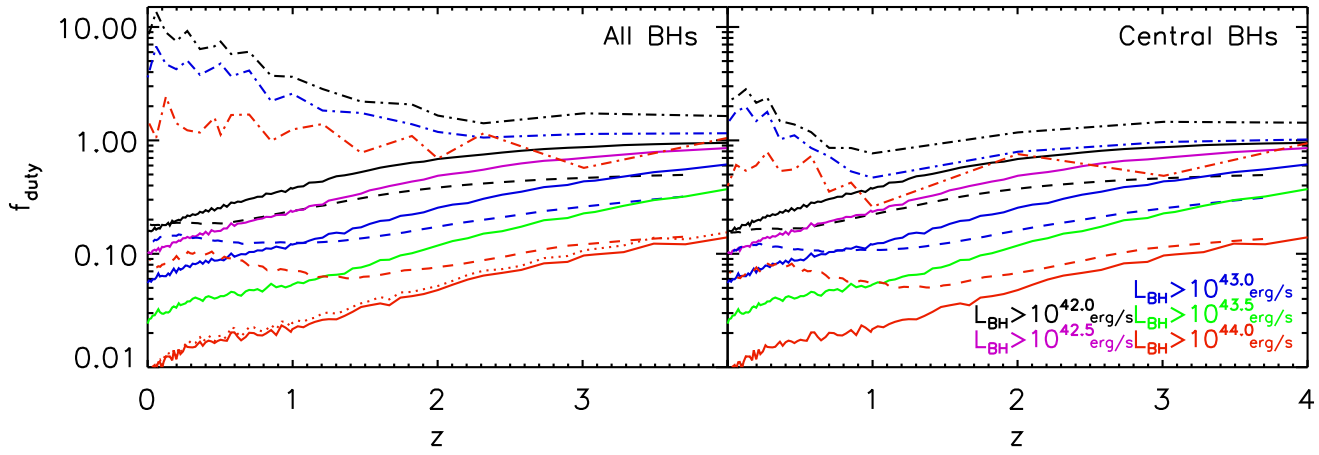

Figure 11. Duty cycle as a function of redshift for varying minimum $L_{\rm BH}$ (line colours, as in Figs 9 and 10). Solid lines: black hole duty cycle for $M_{\rm BH} > 10^6\,{\rm M}_\odot$. Dashed lines: halo duty cycle for haloes above a minimum host mass such that 99 per cent of black holes are included. Dot–dashed lines: bias-predicted duty cycle for all black holes (left), and for central black holes only (right). We find that the black hole duty cycle is virtually identical to the halo duty cycle for $M_{\rm h} > 10^{11.2}\,{\rm M}_\odot$ (see pink dotted line for example), but quite poorly matches the halo duty cycle above a minimum mass determined by either the black hole bias or directly from the distribution of host halo masses.

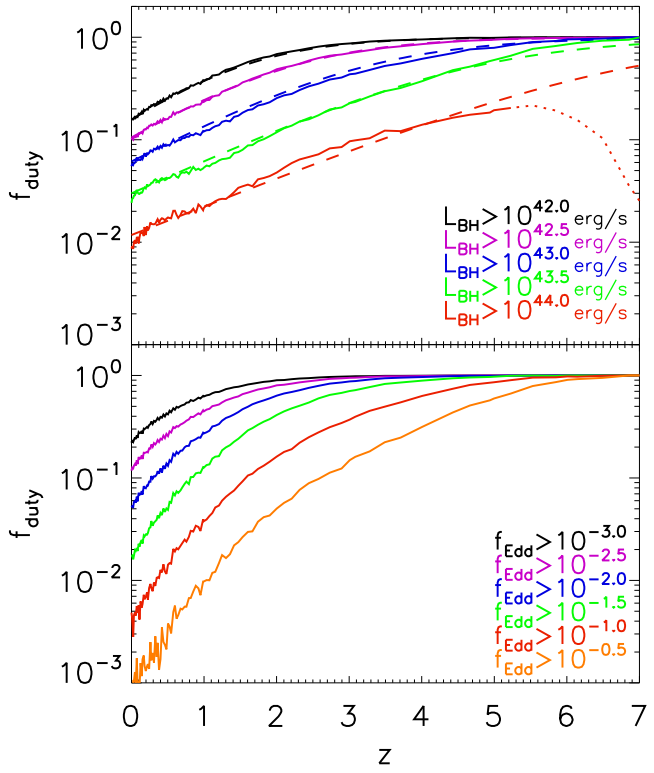

Figure 12. Top: redshift evolution of black hole duty cycle (solid lines; as in Fig. 11), together with our redshift-evolution fit from equation (8) [dashed lines]. Note that the fit is performed for $0 \le z \le 4$, but we plot a larger z-range to show behaviour at higher redshifts. The high-redshift decline for the $L_{\rm BH} > 10^{44}\,{\rm erg\,s^{-1}}$ population is due to the lack of sufficiently massive black holes capable of radiating at such high rates (with the imposed Eddington limit), represented by a dotted line. Bottom: redshift evolution for duty cycles of Eddington fraction-selected black hole populations, which eliminates the high-redshift decline.

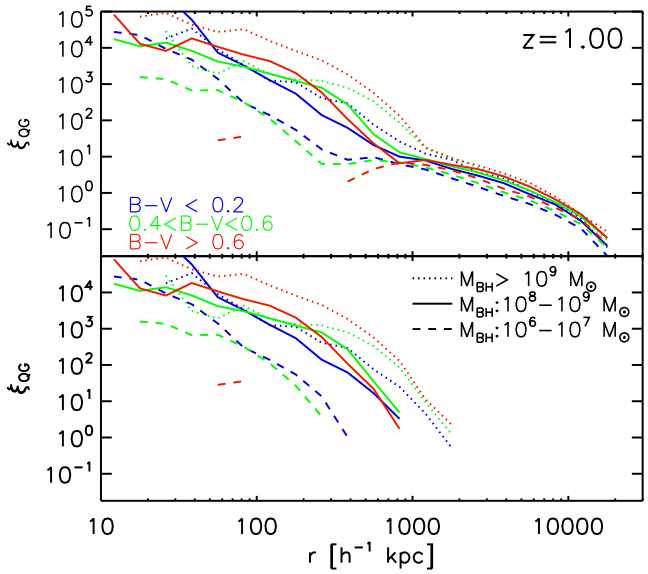

Figure 13. Top: cross-correlation between black holes selected by mass (to characterize integrated black hole feedback) and galaxies selected by $B - V$ colour (to characterize galaxy quenching). Bottom: cross-correlation function 1-halo term only. The top panel shows the cross-correlation of black holes and quenched galaxies is strongly $M_{\rm BH}$ dependent. The lower panel, however, shows that this dependence is actually due to the radial extent of the host halo, which also correlates with $M_{\rm BH}$ (also see Fig. 14).

black holes and galaxies, including black hole mass and luminosity, galaxy mass, stellar luminosity, galaxy colour, star formation rate, etc. The most promising calculation used black holes selected by mass (representing an integrated accretion history) and galaxies selected by colour (characterizing the degree to which star formation has been quenched, while remaining less sensitive to short time-scale variations than specific star formation rate). The top panel of Fig. 13 shows the cross-correlation for these selections, illustrating that the 1-halo clustering signal (below $\sim$1000 h^{-1} kpc) is strongest for massive ($>10^9\,{\rm M}_\odot$) black holes and quenched galaxies ($B - V > 0.6$). In particular, we note that the $B - V < 0.2$ and $0.4 < B - V < 0.6$ behave similar to one another regardless of central black hole mass, while the $B - V > 0.6$ galaxies tend to be more strongly clustered about the most massive black holes, suggesting a possible connection between total energy radiated by the black hole and quenched satellites.

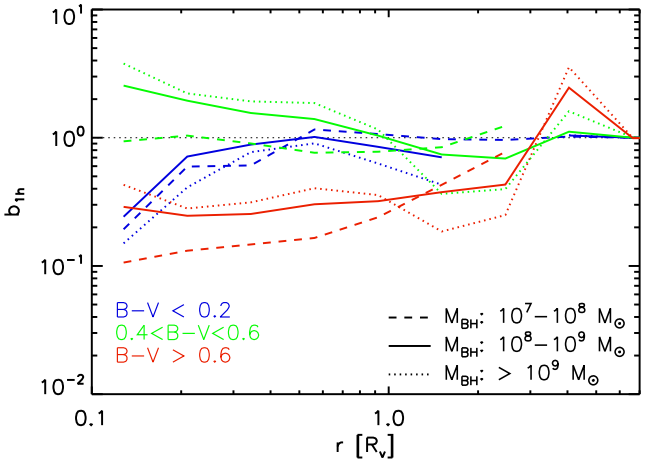

Figure 14. 1-halo bias ($b_{1h} = \sqrt{\xi_{QG}/\xi_{GG}}$) between black hole–galaxy cross-correlation (ξ_{QG}) and galaxy autocorrelation (ξ_{GG}), as a function of host virial radius. Demonstrates no M_{BH} dependence in cross-correlation with quenched galaxies, when controlling for host halo size.

However, further investigation shows that this is not a causal connection between massive black holes and quenched satellite galaxies, but rather a signature of halo radius. In the bottom panel of Fig. 13, we show the 1-halo term only, demonstrating that the difference appears to be largely due to a rescaling of radial separation, with massive black holes tending to be found in galaxies with the largest radial extent. We take this one step further in Fig. 14, showing a bias between the cross-correlation function (ξ_{QG}) and the galaxy–galaxy autocorrelation function (ξ_{GG}; bias defined as $\sqrt{\xi_{QG}/\xi_{GG}}$), with separation defined in units of the virial radius rather than physical size. Here we see that, after rescaling based on the host virial radius, there is no significant difference between any samples. We also considered a possible dependence on the mass of black hole relative to its host halo, and also checked smaller volume simulations with both stronger and weaker radio-mode feedback (as any causal link should be more apparent when feedback is stronger), and confirmed the lack of any quenching signature. This suggests that massive central black holes, although capable of quenching their host galaxies (see, e.g. Sijacki et al. 2015), do not tend to quench star formation of satellite galaxies in the same host halo.

4 CONCLUSIONS

In this work, we have used the Illustris simulation to investigate the clustering of supermassive black holes across a range of redshifts and luminosities. In addition to general agreement with observations, we use the clustering information to link black hole properties to the host masses as well as the AGN duty cycle of black holes and galaxies in general. Our main conclusions are the following:

(i) AGN clustering is found to be luminosity dependent, but primarily at small (1-halo) scales. At larger scales, luminosity dependence primarily occurs at intermediate redshift, where black hole accretion tends to be strongest.

(ii) Correlation length (r_0) can have significant luminosity dependence, especially at intermediate redshifts and when satellite black holes are included. r_0 reaches a minimum at $z \sim 1.5$–2, with higher redshift evolution being strongest for low-luminosity thresholds. Our r_0 estimates are generally lower than observational measures. This is largely due to the limited luminosity range in our simulation (imposed by the simulation volume); however, adjusting observations to match our mean luminosities produces fully consistent results.

(iii) Our estimated black hole bias matches observations very well at low redshift. For $z > 2$, we predict a lower bias than Croom et al. (2005) and Shen et al. (2009), but consistent with Eftekharzadeh et al. (2015).

(iv) AGN clustering tends to be stronger than the expected clustering of haloes of comparable mass; as a result, AGN hosts tend to be less massive than predictions made based on AGN clustering. Although strongest when satellite haloes (found most commonly in the largest haloes) are included, this effect remains even when only central black holes are considered. This suggests that typical host halo masses found based on clustering behaviour may be underestimated by a factor of ∼2, especially at intermediate redshifts.

(v) The scatter in black hole–host scaling relations and typical black hole Eddington fractions results in a wide distribution of host halo masses. Although the distribution for any given $L_{BH,\,min}$ does drop off at low halo mass, there does remain a low-mass tail, especially at low redshifts.

(vi) Due to AGN being more strongly clustered than haloes matched to the typical hosts and both the wide range and low-end tail of the host halo distribution, estimating the minimum host halo mass from AGN clustering tends to substantially overestimate $M_{h,\,min}$, which can have a strong impact on duty cycle estimates.

(vii) At low redshift, the black hole duty cycle follows a power law in M_{BH}, with a normalization set by the luminosity threshold. Higher redshifts also tend to follow a rough power law for $f_{duty} < 0.5$, above which the curve flattens out.

(viii) Black hole duty cycle decreases with time, well fit by a logistic function with lower $L_{BH,\,min}$ thresholds decreasing more rapidly and at lower redshifts.

(ix) Black hole duty cycle is well matched by the halo duty cycle for haloes with $M_h > 10^{11.2}\,M_\odot$, representing a characteristic minimum mass for black hole occupation.

(x) Estimating the duty cycle from AGN number and expected halo number above a given $M_{h,\,min}$ is very inaccurate. In addition to the misestimate of $M_{h,\,min}$, the rapid growth of typical host halo masses at low redshift produces a significant increase in the calculation of f_{duty} which is not found in the true black hole duty cycle.

(xi) We used the AGN–galaxy cross-correlation function to look for a possible signature of AGN-induced quenching of satellite galaxies. Although ξ_{QG} does show M_{BH}-dependent clustering of quenched galaxies, we find this signal is caused by the larger physical size of haloes hosting massive black holes rather than a direct causal link. After controlling for halo size, we find no evidence for AGN inducing quenching in satellite galaxies.

Using the Illustris simulation, we have shown black hole and AGN clustering consistent with current observations, and characterized the luminosity dependence of AGN clustering, which is strongest at intermediate redshift ($z \sim 1.5$–2). One of the most important aspects of clustering analysis is the use of a clustering signal to characterize properties of the host haloes, particularly the halo mass. We find that the typical approach taken (matching AGN clustering to analytic estimates for halo clustering) does very well at high redshift, but can overestimate host mass by ∼50 per cent at low redshift, as low-redshift AGN are found to cluster more strongly than an equivalent-mass halo. Finally, we considered the use of AGN clustering as an estimator for black hole duty cycles. A typical method for this estimation is to assume a minimum host mass for a given AGN luminosity (see equation 3) and a constant duty cycle among haloes above this threshold. Contrary to this

assumption, however, we find a wide distribution of halo masses, including a low-mass tail. This scatter among host masses (as also found in other simulations) must be accounted for or the AGN duty cycle can be strongly overestimated, particularly at low redshift. Overall, we find the black hole duty cycle to evolve smoothly with redshift, and we provide numerical fits characterizing this evolution as well as the dependence on black hole mass and AGN luminosity.

In summary, our work highlights that while black hole clustering is a powerful probe of host halo properties, cosmological simulations, such as Illustris, are needed to fully characterize and account for a number of biases which would otherwise lead to systematically overestimated clustering-predicted host halo masses and black hole duty cycles.

ACKNOWLEDGEMENTS

The authors would like to thank Martin Haehnelt, Volker Springel and Lars Hernquist for their useful comments on this work, and the referee for a very constructive report. CD and DS acknowledge support by the ERC starting grant 638707 'Black holes and their host galaxies: co-evolution across cosmic time'. DS further acknowledges support from the STFC. Simulations were run on the Harvard Odyssey and CfA/ITC clusters, the Ranger and Stampede supercomputers at the Texas Advanced Computing Center as part of XSEDE, the Kraken supercomputer at Oak Ridge National Laboratory as part of XSEDE, the CURIE supercomputer at CEA/France as part of PRACE project RA0844 and the SuperMUC computer at the Leibniz Computing Center, as part of project pr85je.

REFERENCES

Bauer A., Springel V., 2012, MNRAS, 423, 2558
Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289
Bondi H., 1952, MNRAS, 112, 195
Bondi H., Hoyle F., 1944, MNRAS, 104, 273
Bonoli S., Marulli F., Springel V., White S. D. M., Branchini E., Moscardini L., 2009, MNRAS, 396, 423
Bromm V., Larson R. B., 2004, ARA&A, 42, 79
Bromm V., Loeb A., 2003, ApJ, 596, 34
Chabrier G., 2003, PASP, 115, 763
Chatterjee S., Degraf C., Richardson J., Zheng Z., Nagai D., Di Matteo T., 2012, MNRAS, 419, 2657
Croom S. M. et al., 2005, MNRAS, 356, 415
Croton D. J., 2009, MNRAS, 394, 1109
da Ângela J. et al., 2008, MNRAS, 383, 565
DeGraf C., Di Matteo T., Springel V., 2010, MNRAS, 402, 1927
Degraf C., Di Matteo T., Springel V., 2011a, MNRAS, 413, 1383
Degraf C., Oborski M., Di Matteo T., Chatterjee S., Nagai D., Richardson J., Zheng Z., 2011b, MNRAS, 416, 1591
DeGraf C., Di Matteo T., Khandai N., Croft R., Lopez J., Springel V., 2012, MNRAS, 424, 1892
DeGraf C., Dekel A., Gabor J., Bournaud F., 2016, MNRAS, in press
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Eftekharzadeh S. et al., 2015, MNRAS, 453, 2779
Fakhouri O., Ma C.-P., Boylan-Kolchin M., 2010, MNRAS, 406, 2267
Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, ApJ, 703, 1416
Ferrarese L., 2002, ApJ, 578, 90
Gabor J. M., Bournaud F., 2014, MNRAS, 441, 1615
Gebhardt K. et al., 2000, ApJ, 539, L13
Genel S. et al., 2014, MNRAS, 445, 175
Graham A. W., Erwin P., Caon N., Trujillo I., 2001, ApJ, 563, L11
Grazian A., Negrello M., Moscardini L., Cristiani S., Haehnelt M. G., Matarrese S., Omizzolo A., Vanzella E., 2004, AJ, 127, 592
Gültekin K. et al., 2009, ApJ, 698, 198
Haiman Z., Hui L., 2001, ApJ, 547, 27
Häring N., Rix H.-W., 2004, ApJ, 604, L89
Hinshaw G. et al., 2013, ApJS, 208, 19
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2005a, ApJ, 630, 716
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2005b, ApJ, 632, 81
Hopkins P. F., Richards G. T., Hernquist L., 2007, ApJ, 654, 731
Ikeda H. et al., 2015, ApJ, 809, 138
Kereš D., Vogelsberger M., Sijacki D., Springel V., Hernquist L., 2012, MNRAS, 425, 2027
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Kormendy J., Richstone D., 1995, ARA&A, 33, 581
Krolewski A. G., Eisenstein D. J., 2015, ApJ, 803, 4
La Franca F., Andreani P., Cristiani S., 1998, ApJ, 497, 529
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Lidz A., Hopkins P. F., Cox T. J., Hernquist L., Robertson B., 2006, ApJ, 641, 41
McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184
Magorrian J. et al., 1998, AJ, 115, 2285
Martini P., Weinberg D. H., 2001, ApJ, 547, 12
Myers A. D. et al., 2006, ApJ, 638, 622
Myers A. D., Brunner R. J., Nichol R. C., Richards G. T., Schneider D. P., Bahcall N. A., 2007, ApJ, 658, 85
Nelson D., Vogelsberger M., Genel S., Sijacki D., Kereš D., Springel V., Hernquist L., 2013, MNRAS, 429, 3353
Nelson D. et al., 2015, Astron. Comput., 13, 12
Okamoto T., Frenk C. S., Jenkins A., Theuns T., 2010, MNRAS, 406, 208
Oppenheimer B. D., Davé R., 2008, MNRAS, 387, 577
Porciani C., Norberg P., 2006, MNRAS, 371, 1824
Porciani C., Magliocchetti M., Norberg P., 2004, MNRAS, 355, 1010
Puchwein E., Springel V., 2013, MNRAS, 428, 2966
Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013, MNRAS, 430, 2427
Ross N. P. et al., 2009, ApJ, 697, 1634
Shankar F., Crocce M., Miralda-Escudé J., Fosalba P., Weinberg D. H., 2010, ApJ, 718, 231
Shen Y. et al., 2007, AJ, 133, 2222
Shen Y. et al., 2009, ApJ, 697, 1656
Sijacki D., Springel V., di Matteo T., Hernquist L., 2007, MNRAS, 380, 877
Sijacki D., Vogelsberger M., Kereš D., Springel V., Hernquist L., 2012, MNRAS, 424, 2999
Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575
Springel V., 2010, MNRAS, 401, 791
Springel V., 2011, in van de Weijgaert R., Vegter G., Ritzerveld J., Icke V., eds, Invited Review for the Volume "Tesselations in the ciences: Virtues, Techniques and Applications of Geometric Tilings". Springer-Verlag, Berlin, preprint (arXiv:1109.2218)
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V. et al., 2005, Nature, 435, 629
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
Torrey P., Vogelsberger M., Sijacki D., Springel V., Hernquist L., 2012, MNRAS, 427, 2224
Tremaine S. et al., 2002, ApJ, 574, 740
Vogelsberger M., Sijacki D., Kereš D., Springel V., Hernquist L., 2012, MNRAS, 425, 3024
Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031
Vogelsberger M. et al., 2014a, MNRAS, 444, 1518
Vogelsberger M. et al., 2014b, Nature, 509, 177
White M. et al., 2012, MNRAS, 424, 933
Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, MNRAS, 399, 574
Yoshida N., Omukai K., Hernquist L., Abel T., 2006, ApJ, 652, 6

This paper has been typeset from a TEX/LATEX file prepared by the author.